Robust Self-Learning Physical Layer Abstraction Utilizing Optical Performance Monitoring and Markov Chain Monte Carlo

Fanchao Meng, Shuangyi Yan, Rui Wang, Yanni Ou, Yu Bi, Reza Nejabati, Dimitra Simeonidou

High Performance Networks Group, University of Bristol, Tyndall Avenue, Bristol, BS8 1TH, UK
f.meng@bristol.ac.uk

Abstract *We model and experimentally demonstrate a self-learning abstraction process based on statistical assessment of the real-time monitoring data, both amplifier and non-linear noise parameters are periodically updated which further enables an accurate QoT estimator.*

Introduction

The next generation of 5G network is expected to come into standard in the early 2020s, bringing large challenges to the backbone optical networks in terms of capacity, scalability, reliability, etc. Since the traffic from all the other domains finally converges in the packet layer and translates into optical layer traffic demands, accurate optical network resource abstraction is treated as the critical part in provisioning such a reliable inter-network connectivity and programmability. The physical layer abstraction consists of all the parameters needed for control plane metrics computation that is totally separated from the underlying hardware.

On the other hand, as the traffic demands keep increasing to the terabits-per-second era, optical network is likely to reach its Shannon capacity limit. To efficiently utilize the scarce bandwidth resource and further decrease the number of regenerators (CAPEX), optical transparent reach or QoT (Quality of Transmission) margin saving becomes the first consideration. Current vendors leave significant margins to ensure network reliability which results in large reduction in network capacity. These margins can be reduced through careful network planning, where the QoT estimation process plays a key role. The performance of QoT prediction is largely dependent on the mathematical model and its abstraction inputs[1]. Moreover, current and next generation optical networks are fully dynamic and re-configurable where physical layer impairments accumulate along the lightpath that are very hard to be evaluated from theoretical models (such as fiber cuts, nonlinear effects, etc.). In this case a “self-learning” process is required in which the transmission performance is fully observable and the abstractions are self-regulated accordingly making it easier for network debugging and to further enhance the QoT estimator.

In this paper, we propose and experimentally demonstrate a self-learning network based on performance monitoring and Markov Chain Monte Carlo (MCMC) to classify and fit in uncertain physical layer parameters during network planning. A Q-factor estimation error of 0.5dB is achieved through the learning process with well-trained parameters.

Learning Model

Fig. 1(a) shows the overall learning architecture in which control and data plane each has its own planning process while sharing some identical information. It is further separated into three blocks: abstraction inputs, system modelling and observations. Information (abstraction inputs) such as data rates, source & destination pairs, etc. are treated as control plane inputs. Topology, frequency, power, etc. are inputs that are shared by both control and data planes. The data plane inputs are the parameters of interest which are (a) nonlinear distortion coefficient α_{NL}, (b) filter wavelength detuning $\Delta\lambda_0$, (c) amplifier gain G(G>>1), (d) EDFA (Erbium-doped Fiber Amplifier) noise figure *NF*, any other parameters influencing the QoT performance are neglected[2]. The inputs are fed into a model (system modelling) based on Eq. (1)(2)(3)[3] to estimate link and path performance, where P_{in} is the output power per channel of each EDFA, P_{ASE} is the noise power density, *h* is Planck’s constant, *v* is the optical carrier frequency, B=12.48GHz.

$$\mathrm{OSNR}_{link} = P_{in}/(P_{ASE} + \alpha_{NL} * P_{in}^3) \quad (1)$$

$$P_{ASE} = h * v * NF * G * B \quad (2)$$

$$OSNR_{path} = (\textstyle\sum_{k=1}^{n} OSNR_{Link\#k}^{-1})^{-1} \quad (3)$$

A calibration parameter is used to align the model output data to the monitoring data first. The performance value such as Q factor calculated based on OSNR is further uploaded to the controller for RSWA (Routing, Spectrum and Wavelength Assignment) computation. As long as physical layer services are provisioned and monitored(observations), input parameter tuning is performed periodically until the estimation and monitoring values converge. We separate some of the parameters first as they have compound

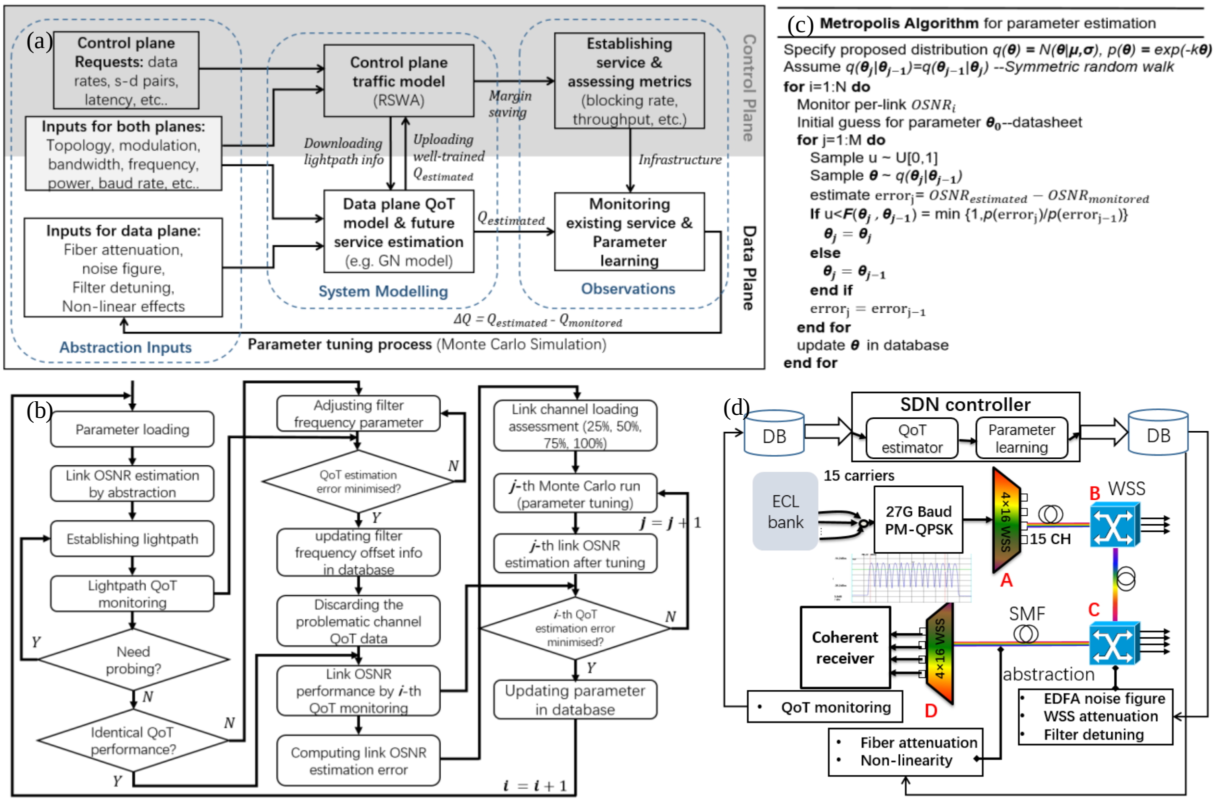

Fig. 1: (a) overall architecture of the learning model ("s-d" stands for source-to-destination), (b) flow chart of parameter learning process, (c) Metropolis Algorithm for sampling parameters $\boldsymbol{\alpha_{NL}}$, ***G*** and ***NF,*** (d) network testbed setup ("DB" for database).

influence on the received signal quality. The flow chart of parameter separation process breaks down into several steps in Fig. 1(b): <1>parameter loading & network planning; <2>in the case channels sharing identical lightpath present different reception performance, the worse channel is more likely to suffer from filter wavelength detuning. (active probing is triggered if not enough channel to compare); <3>MCMC is triggered for $\Delta\ \boldsymbol{\lambda_0}$ sampling until lightpath Q estimation error is minimized; <4>computing per-link OSNR based on per-path metrics from $OSNR_{link\#k} = (OSNR_{path\#k}^{-1} - OSNR_{path\#(k-1)}^{-1})^{-1}$, $OSNR_{path\#(k-1)}^{-1}$ is derived from the regression line of other wavelengths through supervised learning; <5>computing link estimation error ΔQ_{link}; <6> link channel loading assessment[4], in this experiment there are four cases (single channel, 4 channels, 8 channels, 15 channels) where non-linear distortion coefficient α_{NL} is estimated using Gaussian Noise (GN) model[3]. α_{NL} has upper and lower bound for example in the case of 6 channels, $\alpha_{NL-4\ ch} < \alpha_{NL-6\ ch} < \alpha_{NL-8\ ch}$; <7>MCMC is triggered for α_{NL}, *G* and *NF* tuning until $\Delta\ Q_{link}$ is minimized; <8>updating abstraction database and return to <1>.

Fig. 1(c) shows the pseudocode of MCMC process in which Metropolis Algorithm (MA) is applied to reduce the parameter uncertainty. In the case of step <7>, we investigate the compound value of *c=G*NF* of interest as each wavelength in each span undergoes identical fiber attenuation and EDFA NF. The initial samples for *G, NF* and α_{NL} come from datasheet and model simulation (maximum likelihood). The distribution of potential *G', NF'*(or *c'*) and α_{NL}' for the succeeding sampling follows a Gaussian distribution *q(θ) = N(θ|μ,σ)* (symmetric random walk) which is centered at previous accepted value *μ* so that the new guess is only dependent on the previous guess (Markov chain) as P($\theta_n|\theta_{n-1},\ldots,\theta_2,\theta_1$)=P($\theta_n|\theta_{n-1}$). A likelihood function which is maximized by the QoT model output exactly fitting the monitoring data is chosen as *p(x) = exp(-kx)* where *x* is the estimation error for each MCMC. A new sample is accepted if the estimation error $\Delta\ Q_{link}$ decreases, otherwise there is a probability of 1-$p(\mathrm{error_j})/p(\mathrm{error_{j-1}})$ that it (j-th guess) is rejected.

Testbed setup and results

Fig. 1(d) shows the testbed network consisting of a coherent transponder (A and D) supporting up to 15 equalized 50GHz-spaced 27Gbaud DP-QPSK transmission and two intermediate nodes

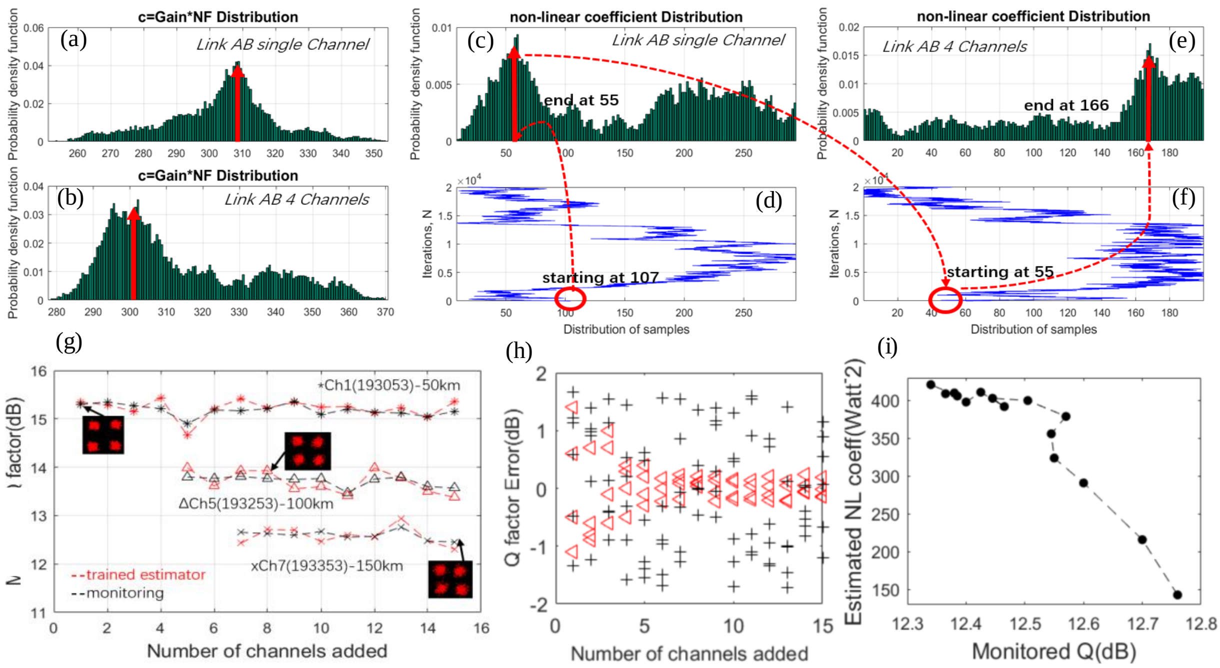

Fig. 2: (a) pdf (probability density function) of 1-channel $\boldsymbol{c}$ distribution, (b) pdf of 4-channel $\boldsymbol{c}$ distribution, (c) pdf of 1-channel $\boldsymbol{\alpha_{NL}}$ distribution, (d)1-channel $\boldsymbol{\alpha_{NL}}$ iteration path, (e) pdf of 4-channel $\boldsymbol{\alpha_{NL}}$ distribution, (f) 4-channel $\boldsymbol{\alpha_{NL}}$ iteration path, (g) monitored & estimated Q-factor with channel loading, (h) estimation error with(red) & without(black) machine learning, (i) estimated $\boldsymbol{\alpha_{NL}}$ vs Q.

(B and C, emulated by 4×16WSS) which drop 3 and 4 wavelengths respectively for coherent detection. Each node is connected by 50km 0.2 dB/km SMF, the WSS insertion loss is 6dB, all devices are cross-connected to a 192×192 optical switch (Polatis) which gives 1dB loss in each port. EDFA compensates all the span loss giving a launch power of +1dBm/channel. Fig. 2(a) and (b) shows the result of MCMC sampling (20000 iterations each time) pdf for parameter *c* in the case of one channel and 4 channels respectively. The actual parameter is most likely to be where the density is the highest (marked by red arrows), *c* stays at the value 300-310 which means the noise figure and fiber attenuation are not impacted by channel loading. Fig. 2(c) and (e) shows the pdf for α_{NL} which changes as the channels are loaded. The pdf of α_{NL} has a flatter region due to its less impact to OSNR compared to *c*. The iteration starts from α_{NL} =107(from theoretical model) as seen in Fig. 2(d) and ends at 55, after three more channels are loaded simultaneously (Fig. 2(f)), the iteration starts at 55 and ends at 168. Fig. 2(g) shows the Q-factor estimation with well-trained parameters for three channels (Ch1, Ch5, Ch7) that are terminated after 50km, 100km 150km respectively. As the number of channels increases, Q-factor slightly decreases due to the added non-linear distortion. The estimated value fits the monitoring data very well with a standard deviation of 0.39dB and maximum estimation error of 0.5dB. Fig. 2(h) shows the estimation error with and without learning, the estimation error is reduced to ±0.24dB as more data (channels) are added and monitored. Fig 2(i) shows the worse case $\boldsymbol{\alpha_{NL}}$ learning against monitored Q-factor after 150km, as Q decreases after 12.5dB because of loading, α_{NL} keeps almost constant because the added channels no longer have significant impact on QoT.

Conclusion

We have experimentally demonstrated a “self-learning” network with dynamic abstraction process based on real-time monitoring and Markov Chain Monte Carlo. The learning process minimizes the abstraction uncertainties resulting in a robust QoT estimator with an error reduction from 3.7dB (without learning) to 0.5dB.

Acknowledgement

This work is supported by EPSRC grant EP/L020009/1 TOUCAN and EP/L026155/2 INSIGHT project.

References

[1] Y. Pointurier et al., “Design of Low-margin Optical Networks,” Proc. OFC, Tu3F.5 (2016).

[2] J. Pesic et al., “Faster return of investment in WDM networks when elastic transponders dynamically fit ageing of link margins,” Proc. OFC, M3K.2 (2016).

[3] P. Poggiolini, "The GN Model of Non-Linear Propagation in Uncompensated Coherent Optical Systems," JLT, vol. 30, pp. 3857-3879, 2012.

[4] R. Wang, “Load and Nonlinearity Aware Resource Allocation in EON,” Proc. OFC, W1H.5 (2017).